AMENDMENT TO THE CLAIMS

1. (Original) A method of ascertaining phoneme speech unit boundaries of adjacent speech units in speech data, the method comprising:

receiving training data of speech waveforms with known boundary locations of phoneme speech units contained therein;

processing the speech waveforms to obtain multi-frame acoustic feature pseudo-triphone representations of a plurality of pseudo-triphones in the speech data, each pseudo-triphone comprising a boundary location, a first phoneme speech unit preceding the boundary location and a second phoneme speech unit following the boundary location;

clustering the multi-frame acoustic feature pseudo-triphone representations as a function of acoustic similarity in a plurality of clusters;

training a refining model for each cluster;

receiving a second set of data of speech waveforms with initial boundary locations of adjacent phoneme speech units contained therein;

identifying pseudo-triphones in the second set of data and corresponding refining models for each of the pseudo-triphones; and

using the refining model for each corresponding pseudo-triphone to locate a boundary location different than the initial boundary.

2. (Original) The method of claim 1 wherein clustering comprises maintaining a minimum number of multi-frame acoustic feature pseudo-triphone representations greater than one in each cluster.

3. (Original) The method of claim 1 wherein clustering comprises controlling a number of clusters created.

4. (Original) The method of claim 1 wherein clustering comprises using a Classification and Regression Tree clustering technique.

5. (Original) The method of claim 1 wherein clustering comprises using a Support Vector Machine (SVM) clustering technique.

6. (Original) The method of claim 1 wherein clustering comprises using a Neural network (NN) clustering technique.

7. (Original) The method of claim 1 wherein clustering comprises using a vector quantization (VQ) clustering technique.

8. (Original) The method of claim 1 wherein processing the speech waveforms to obtain multi-frame acoustic feature pseudo-triphone representations comprises forming a multi-dimensional matrix or vector based on a number of frames of speech waveform data adjacent to the known boundary.

9. (Original) The method of claim 8 wherein forming a multi-dimensional matrix or vector comprises reducing the number of dimensions.

10. (Original) The method of claim 1 wherein training a refining model for each cluster comprises forming a Gaussian Mixture Model to model the most likely locations of a boundary for each cluster.

11. (Original) The method of claim 10 wherein forming a Gaussian Mixture Model to model the most likely locations of a boundary

for each cluster comprises obtaining only a single Gaussian component.

12. (Original) The method of claim 1 wherein training a refining model for each cluster comprises forming a Neural Network model to model the most likely locations of a boundary for each cluster.

13. (Original) The method of claim 1 wherein training a refining model for each cluster comprises forming a Hidden Markov Model to model the most likely locations of a boundary for each cluster.

14. (Original) The method of claim 1 wherein training a refining model for each cluster comprises forming a Maximum Likelihood Probability model to model the most likely locations of a boundary for each cluster.

15. (Cancelled)
16. (Cancelled)
17. (Cancelled)
18. (Cancelled)
19. (Cancelled)
20. (Cancelled)
21. (Cancelled)
22. (Cancelled)
23. (Cancelled)
24. (Cancelled)
25. (Cancelled)
26. (Cancelled)
27. (Cancelled)
28. (Cancelled)
29. (Cancelled)

30. (Cancelled)
31. (Cancelled)
32. (Cancelled)
33. (Cancelled)
34. (Cancelled)